Electronic Supplementary Information

Theophylline-induced synergic activation of guide RNA to control CRISPR/Cas9 function

Yan Liu[a‡], Yang Wang[a‡], Jiao Lin[a], Liang Xu[a]*

[a]MOE Key Laboratory of Bioinorganic and Synthetic Chemistry, School of Chemistry, Sun Yat-Sen University, Guangzhou, 510275, China

*Email: xuliang33@mail.sysu.edu.cn

‡These authors contributed equally to this work.

Materials and Methods

Materials

All chemical reagents were purchased from Sangon Biotechnology (Shanghai, China). T7 High Yield Transcription Kit, DNA gel extraction kit, Hiscript III RT SuperMix and ChamQ Universal SYBR qPCR Master Mix were purchased from Vazyme. DNase I, Ni-NTA and Lipofectamine 3000 were purchased from Thermo Fisher. Opti-MEM media and fetal bovine serum (FBS) were purchased from Gibco. All DNA sequences and the FAM-labeled DNA substrate used in this study were ordered from Sangon Biotechnology.

Preparation of gRNA

The gRNA was in vitro transcribed by T7 RNA polymerase using the T7 High Yield Transcription Kit (Vazyme) according to manufacturer's instructions. Briefly, the duplex DNA construct containing the T7 promoter and coding sequence of gRNA was prepared and purified by commercial DNA gel extraction kits. The in vitro transcription

reaction was performed by T7 polymerase and incubated for 12 hr at 37 ℃. After reaction, the DNA template was digested by DNase I, and the reaction mixture was subsequently processed by phenol/chloroform extraction, followed by ethanol precipitation. The RNA product was desalted by the Bio-Spin 6 column (Bio-Rad), and stocked in DEPC-treated water at -80 °C for following experiments.

dCas9 Protein purification

The coding sequence for dCas9 (D10A and H840A) was cloned into a pET28a expression vector to construct pET28a-dCas9 with an N-terminal hexa-histidine (6xHis) tag. The protein was expressed in E. coli strain BL21(DE3) similarly as described. Briefly, BL21(DE3) transformed with pET28a-dCas9 was grown in LB medium supplemented with 50 μg/mL kanamycin at 37 ℃ to reach OD_{550} 0.6. Then, the temperature was cooled down to 16 ℃, and the expression was induced with IPTG (0.25 mM) overnight (~16 hr). The protein was purified by Ni-NTA after lysis of cells in 20 mM Tris pH 7.4, 500 mM NaCl (supplemented with 10 mM β-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride). The purified protein was concentrated to ~20 μM in 20 mM Tris-HCl (pH 7.4) with 500 mM NaCl, 1 mM DTT and 5% Glycerol, flash-frozen in liquid nitrogen, and stored at -80 ℃.

Electrophoretic mobility shift assay (EMSA) for measurement of dCas9 binding ability

To check the binding ability of dCas9, the duplex DNA substrate (see table S2 for sequence information) was labeled by 5′-FAM to monitor the gel shift behavior in the presence of different gRNAs. For evaluation of binding behavior, gRNA (1 μM) was premixed with theophylline before addition of dCas9 protein (1 μM), FAM labeled-dsDNA (100 nM) in the binding buffer (20 mM Tris-HCl, 100 mM KCl, l, 5 mM $MgCl_2$, 1 mM DTT, 0.1 mM EDTA). The mixture was incubated at 37 ℃ for 10 min, and then analyzed by electrophoresis in 5% native polyacrylamide gel on ice, and imaged by Gel Imaging System (Tanon 2500R).

To quantify the binding ability, a single-site binding equation was utilized to fit the

changes of percentages for the complex formation (Eq. 1).

$[Product]=B_{max}*X/(K_d+X)$ (1)

where B_{max} is the maximal binding, and K_d is the concentration of ligand required to reach half-maximal binding. X represents the ligand concentration, and [Product] indicates the change of complex formation compared to the absence of ligand.

Biochemical cleavage assay for measurement of Cas9 nuclease activity

Cleavage assay was performed using a 2-kb PCR amplified DNA substrate. Briefly, 100 nM Cas9, 100 nM gRNA, and 10 ng/μL DNA substrate were incubated with or without theophylline in the cleavage buffer (20 mM Tris-HCl, 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 0.1 mM EDTA) at 37 °C for designated time points. Reactions were quenched by addition of EDTA (0.5 M), followed by treatment of proteinase K at 37 °C for 10 min. The cleaved product was analyzed by electrophoresis in 1% agarose gel, and imaged by Gel Imaging System (Tanon 2500R).

Gel shift assay for gRNA-Cas9 binding

The gRNA was in vitro transcribed by T7 RNA polymerase using the T7 High Yield Transcription Kit as described above. For FAM-labelling of the gRNA, a ssDNA (1 μM) with 3'-FAM modification, paired with the spacer sequence of gRNA, were mixed with the gRNA (1 μM) in 5 mM Tris-HCl (pH 7.4)/10 mM NaCl buffer. Annealing was carried out by heating at 95 °C for 5 min and cooling down to 25 °C at a rate of −1 °C/min to label the gRNA. Given that the spacer region of gRNA must pair with the DNA substrate upon the formation of CRISPR/Cas9 complex, this approach of fluorescence-labeling on gRNA may not significantly influence the binding behaviors between gRNA and the Cas9 protein. Purified dCas9 was mixed with 20 nM of the designated gRNA (see Table S3 for sequence information) and incubated at 37 °C for 10 min. The concentrations of the dCas9 were varied from 0-50 nM. Afterwards, the mixtures were loaded onto 5% native PAGE (in 0.5xTBE running buffer) and the electrophoresis was run at room temperature for 1 hr. The native gels were then visualized by Multifunctional Laser Scanning Imaging System (GE, Typhoon).

Gene activation by CRISPRa in HEK293T cells

First, HEK293T cells were packaged by lentivirus to generate the stably expressed CRISPRa cell line. Briefly, DNA sequence encoding dCas9 (D10A and H840A) with VP64-p65-Rta (VPR) and EGFP fused to its C-terminus was cloned into pHR-SFFV (Addgene #79121) lentiviral vector to generate pHR-SFFV-dCas9-VPR-EGFP for constitutive expression of dCas9-VPR-EGFP. To obtain Viral production cells, 1,000,000 cells were seeded into 6-well plates before co-transfection of 1000 ng pMD2.G (Addgene #12259), 2000 ng pCMV-dR8.2 (Addgene #84550) and 3000 ng pHR-SFFVdCas9-VPR-EGFP using 16 μL P3000 and 7 μL Lipofectamine 3000. The transfection medium was then replaced by 2 mL fresh medium after 6 hr incubation. 24 hr later, the culture medium was collected and filtered to remove cells, and treated with 10 μg/μL 1.6 μL polybrene before infecting the freshly seeded 293T cells with a density of 20, 000 cells in 6-plate well. The infected cells were then collected and sorted by flow cytometer (Beckman moFlo XDP) to obtained 293T-dCas9-VPR cells.

HEK293T cells were cultured in complete media, DMEM (Dulbecco`s modification of Eagle medium, Gibco), 10% (v/v) FBS (fetal bovine serum, Gibco) and 1% penicillin/streptomycin (Invitrogen), at 37 ℃ in a 5% CO_2 incubator. Cells (2.5×10^5 per well) were seeded into 6-well plates, and immediately transfected with 2000 ng plasmids which coding sgRNA targeting the ASCL1 or CXCR4 gene using 5.5 μL Lipofectamine 3000 mixed in 500 μL Opti-MEM media. The negative control (NC) indicated the gRNA could not target the ASCL1 or CXCR4 gene; the positive control (PC) indicated the standard gRNA without introduction of the theophylline aptamer. The medium was then replaced with a complete DMEM after 6 hr incubation, then PBS or 4 mM theophylline were added and celles were cultured for another 18 hr .Total RNA from 293T cells was isolated using trizol, chloroform, isopropanol and ethanol and quantified by the UV–Vis spectrophotometry.

RNA extraction and real-time quantitative PCR

To quantify mRNA expression level, 1 μg total RNA was used to generate cDNA using Hiscript III RT SuperMix for qPCR according to manufacturer’s instructions.

Expression levels of GAPDH and target gene ASCL1 and CXCR4 were detected by CFX96 Real-Time System (Bio-Rad) using ChamQ Universal SYBR qPCR Master Mix with the following PCR cycle: 95 ℃, 10 min; 40 cycles of 95 ℃, 15 s and 61 ℃, 1 min; 65 ℃to 95 ℃, 0.5 ℃/s. GAPDH was used as the internal control, and the data were normalized to the expression of GAPDH. Relative gene expression was calculated using the Delta-Delta-Ct (ddCt) algorithm.

Tables for Sequence Information

Table S1. gRNA sequences used in this study

name	Sequence (5′-3′)
negetive control	GGGAACGACUAGUUAGGCGUGUAGUUUUAGAGCUAGAAAU AGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU
ASCL1 positive control	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCUAGAAAU AGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU
ASCL1-TL-CM 1	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGATACCAGC CGAAAGGCCCTTGGCAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
ASCL1-TL-CM 2	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGACGATACCAG CCGAAAGGCCCTTGGCAGCGAAGUUAAAAUAAGGCUAGUCC GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
ASCL1-TL-CM 3	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAAGATACCAG CCGAAAGGCCCTTGGCAGCUAAGUUAAAAUAAGGCUAGUCC GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
ASCL1-TL-CM 4	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCGATACCA GCCGAAAGGCCCTTGGCAGCGUAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU

ASCL1-TL-CM 5	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGAGATACCAGCCGAAAGGCCCTTGGCAGCUUAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
ASCL1-TL-CM 6	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGGCGATACCAGCCGAAAGGCCCTTGGCAGCGUCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
ASCL1-SL2-CM 1	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAGATACCAGCCGAAAGGCCCTTGGCAGCGGCACCGAGUCGGUGCU
ASCL1-SL2-CM 2	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGATACCAGCCGAAAGGCCCTTGGCAGCGGGCACCGAGUCGGUGCU
ASCL1-SL2-CM 3	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAGATACCAGCCGAAAGGCCCTTGGCAGCUGGCACCGAGUCGGUGCU
ASCL1-SL2-CM 4	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAGCGATACCAGCCGAAAGGCCCTTGGCAGCGUGGCACCGAGUCGGUGCU
ASCL1-SL2-CM 5	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAGAGATACCAGCCGAAAGGCCCTTGGCAGCUUGGCACCGAGUCGGUGCU
ASCL1-SL2-CM 6	GGGAUGGAGAGUUUGCAAGGAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAGGCGATACCAGCCGAAAGGCCCTTGGCAGCGUCGGCACCGAGUCGGUGCU
CXCR4 positive control	GGGACGCGAGGAAGGAGGGCGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
CXCR4-TL-CM 2	GGGACGCGAGGAAGGAGGGCGCGUUUUAGACGATACCAGCCGAAAGGCCCTTGGCAGCGAAGUUAAAAUAAGGCUAGUCCGU

	UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
CXCR4-TL-CM 3	GGGACGCGAGGAAGGAGGGCGCGUUUUAGAAGATACCAGCC GAAAGGCCCTTGGCAGCUAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
CXCR4-SL2-CM 2	GGGACGCGAGGAAGGAGGGCGCGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGATACCAGCC GAAAGGCCCTTGGCAGCGGGCACCGAGUCGGUGCU
CXCR4-SL2-CM 3	GGGACGCGAGGAAGGAGGGCGCGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAGATACCAGCC GAAAGGCCCTTGGCAGCUGGCACCGAGUCGGUGCU

Table S2. DNA sequences used for EMSA

FAM-dsDNA-1	FAM-GGGCCTTGCAGTGGGCGCATGGAGAGTTTGCAAGGAGC **CGG**CGCTTTGGTCGGCATGGC
dsDNA-2	GCCATGCCGACCAAAGCG**CCG**GCTCCTTGCAAACTCTC CATGCGCCCACTGCAAGGCCC

Table S3. gRNA sequences and 3'-FAM labeled ssDNA used for gRNA binding

Standard gRNA	GGGCCAUAGCCAGUAGUCUCCUUGAAGGUUUUAGA GCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU
TL-CM-3 gRNA	GGGCCAUAGCCAGUAGUCUCCUUGAAGGUUUUAGA AGATACCAGCCGAAAGGCCCTTGGCAGCUAAGUUAA AAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCU
SL2-CM-2 gRNA	GGGCCAUAGCCAGUAGUCUCCUUGAAGGUUUUAGA GCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAAGATACCAGCCGAAAGGCCCTTGGCAGCUGG

	CACCGAGUCGGUGCU
3’-FAM labeled ssDNA	CTTCAAGGAGACTACTGGCTATGG-FAM

Table S4. Primers (5'-3') used in qPCR assay in HEK293T cells.

Gene	F-primer	R-primer
GAPDH	ACAGTCAGCCGCATCTTCTT	ACGACCAAATCCGTTGACTC
ASCL1	GGAGCTTCTCGACTTCACCA	AACGCCACTGACAAGAAAGC
CXCR4	CCCTTGGAGTGTGACAGCTT	TTGTGGGTGGTTGTGTTCCA

Supplementary Figures

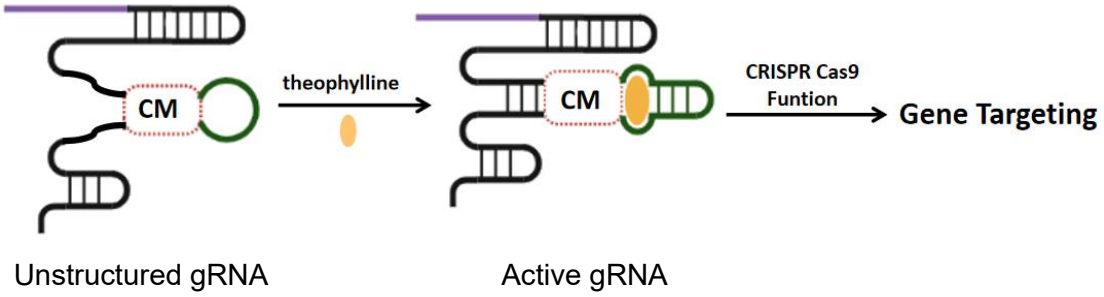

Figure S1. Scheme for the allosteric regulation of gRNA by the theophylline aptamer in stem loop 2 (SL2) hairpin. In the absence of theophylline, the gRNA with the theophylline aptamer was unstructured; treatment of theophylline would induce the folding of the theophylline aptamer and consequently facilitate the formation of active gRNA to achieve the gene targeting function.

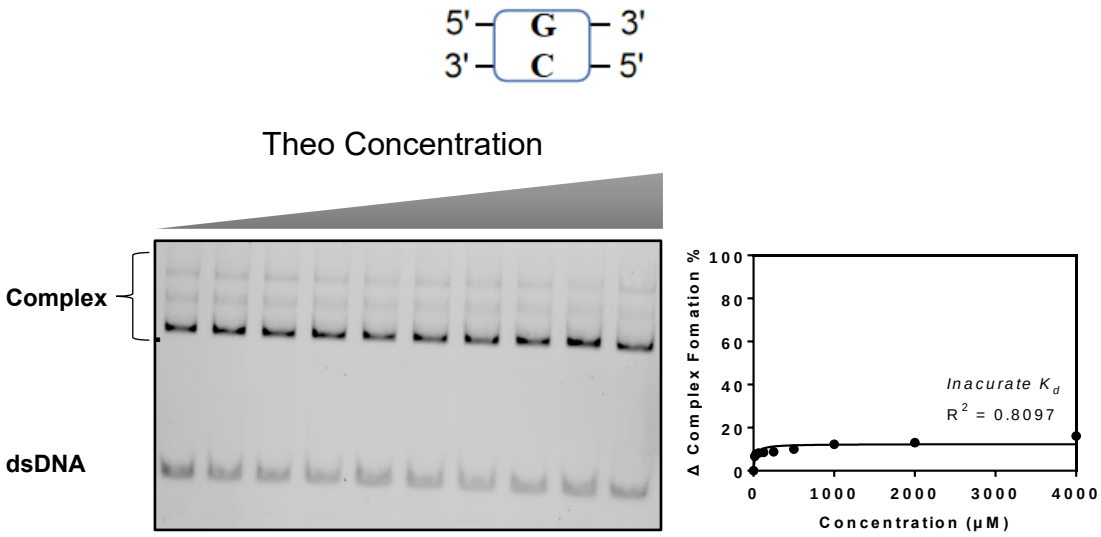

Figure S2. Concentration-dependent regulation of the complex formation with the **TL-CM-1** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively. Due to the weak change of complex formation, the K_d value could not be determined effectively.

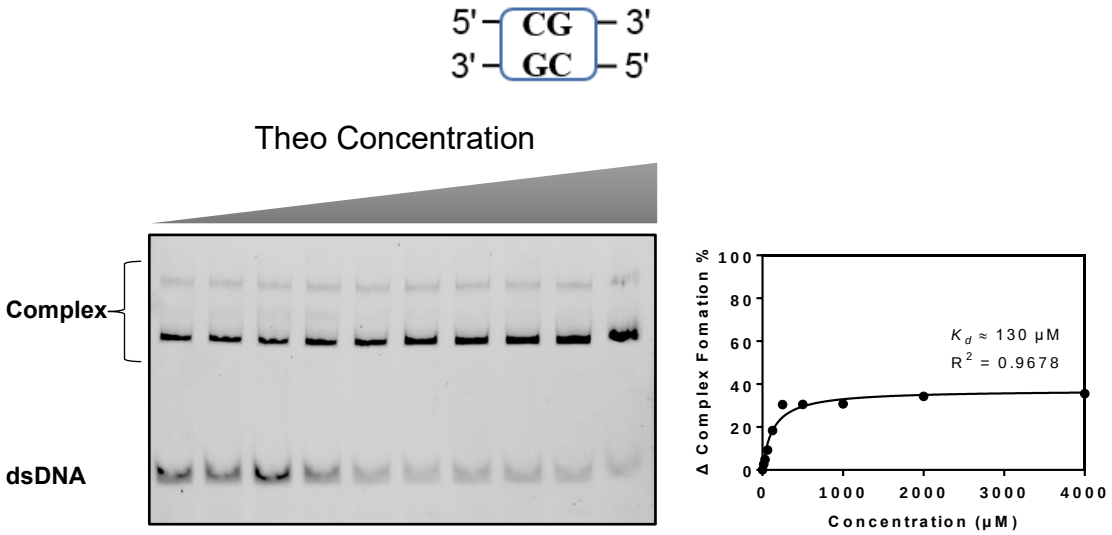

Figure S3. Concentration-dependent regulation of the complex formation with the **TL-CM-2** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

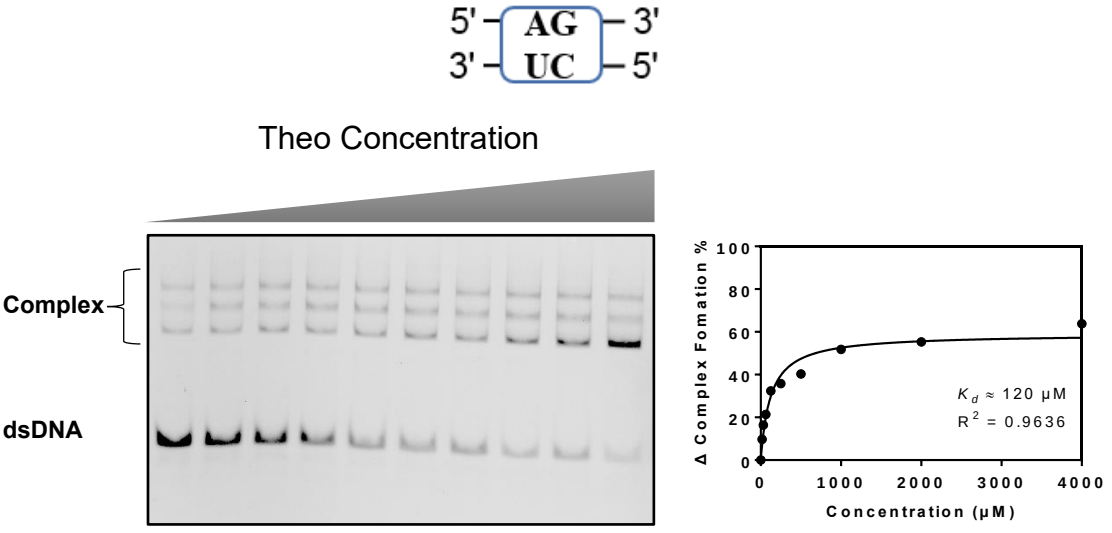

Figure S4. Concentration-dependent regulation of the complex formation with the **TL-CM-3** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

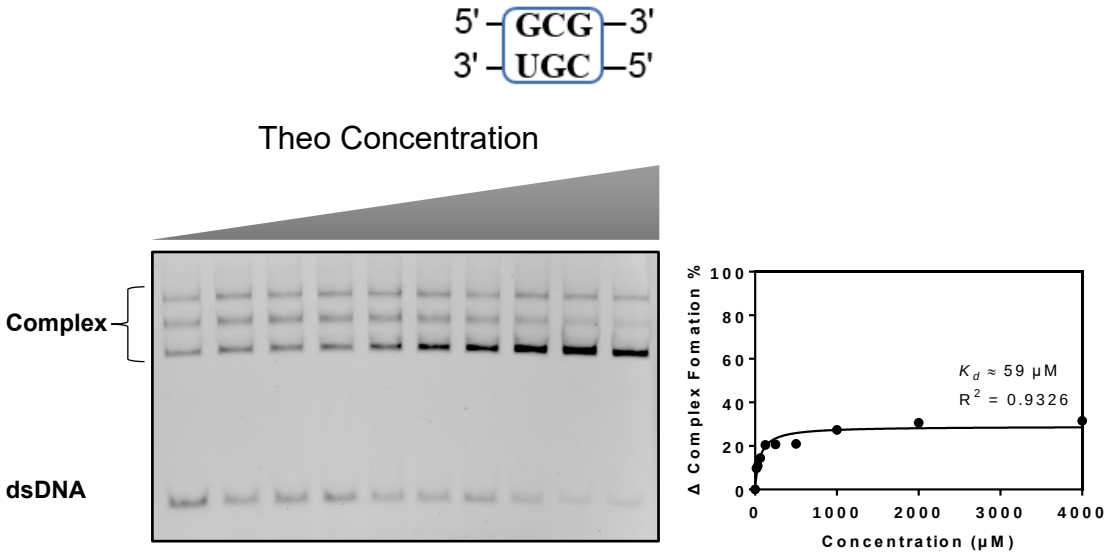

Figure S5. Concentration-dependent regulation of the complex formation with the **TL-CM-4** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

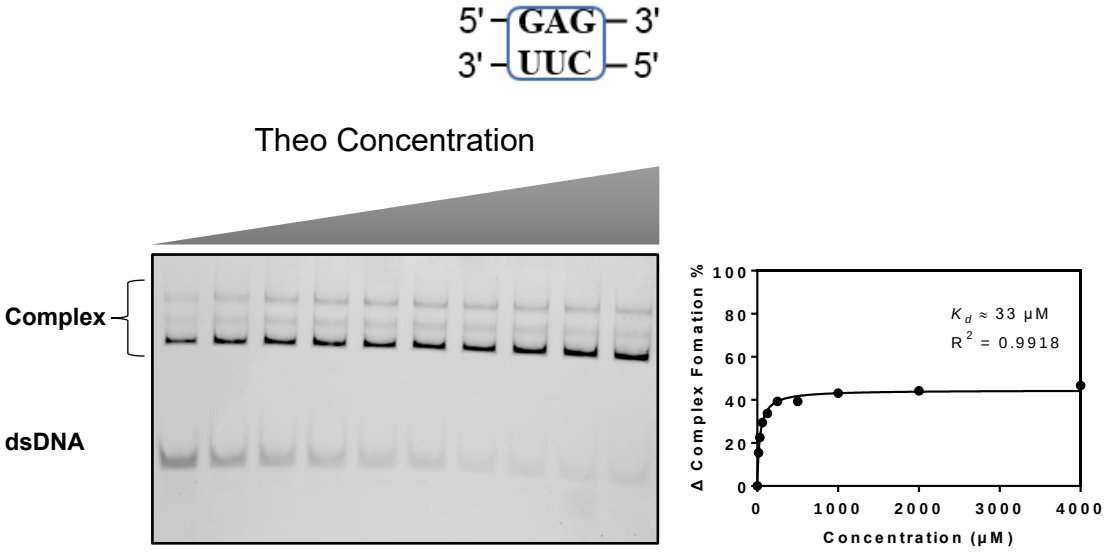

Figure S6. Concentration-dependent regulation of the complex formation with the **TL-CM-5** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

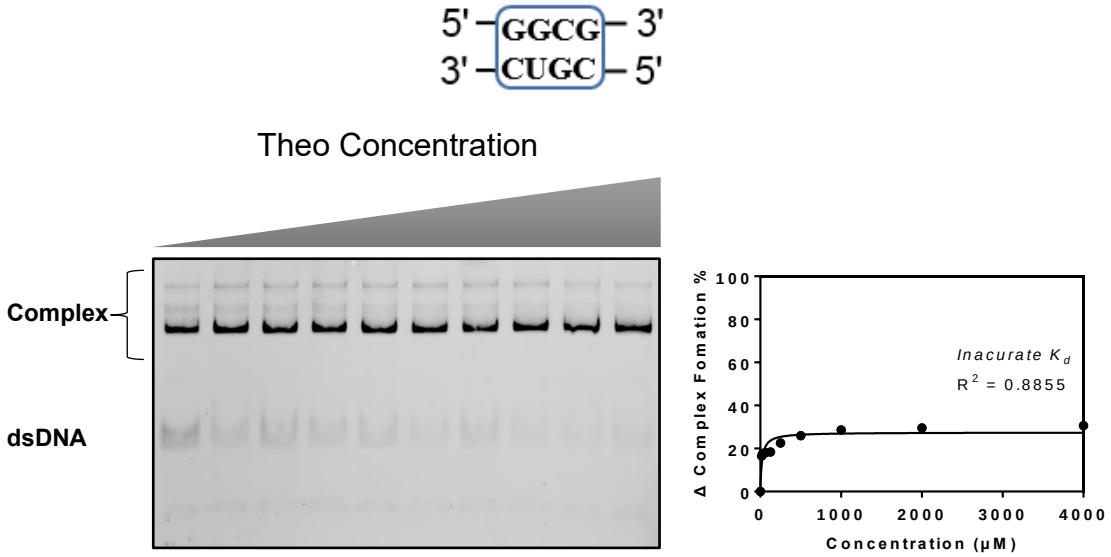

Figure S7. Concentration-dependent regulation of the complex formation with the **TL-CM-6** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively. Due to the weak change of complex formation, the K_d value could not be determined effectively.

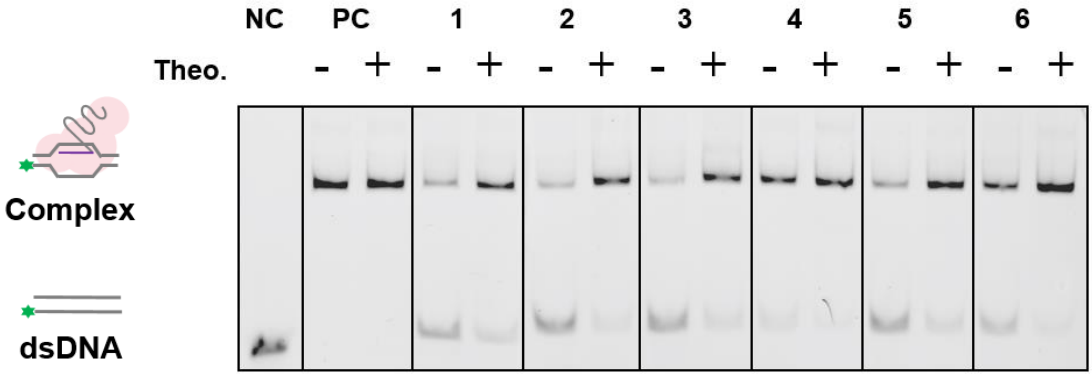

Figure S8. Theophylline-regulated binding abilities of the CRISPR/Cas9 complex with engineering of the stem loop 2 hairpin (the Aptamer@SL2 design). The gel shift assay indicated theophylline-induced complex formation with different designs of CMs. The numbers indicated the designs of CMs as described in Figure 2a. The negative control (NC) indicated the targeting duplex DNA without any gRNA; the positive control indicated formation of the standard gRNA/dCas9/dsDNA ternary complex without introduction of the theophylline aptamer. The concentration of theophylline (Theo) was either 0 (-) or 4 mM (+) for each lane. The sequence designs of CMs were listed.

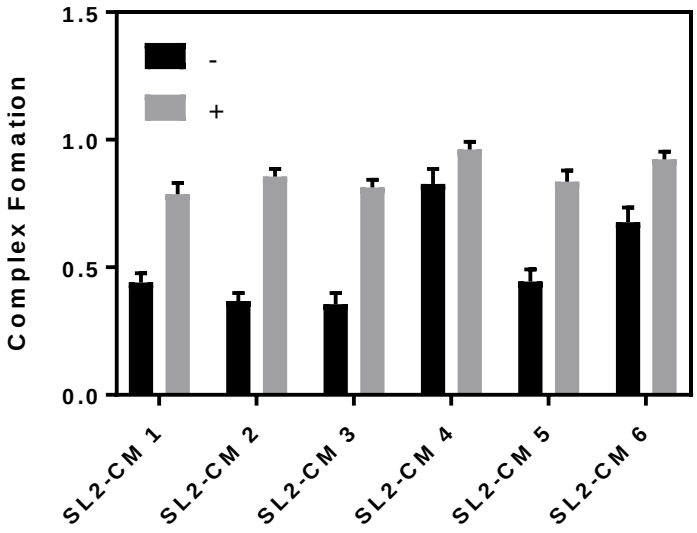

Figure S9. Quantitative analysis of the complex formation before and after treatment of 4 mM theophylline with different CMs based on the Aptamer@SL2 design.

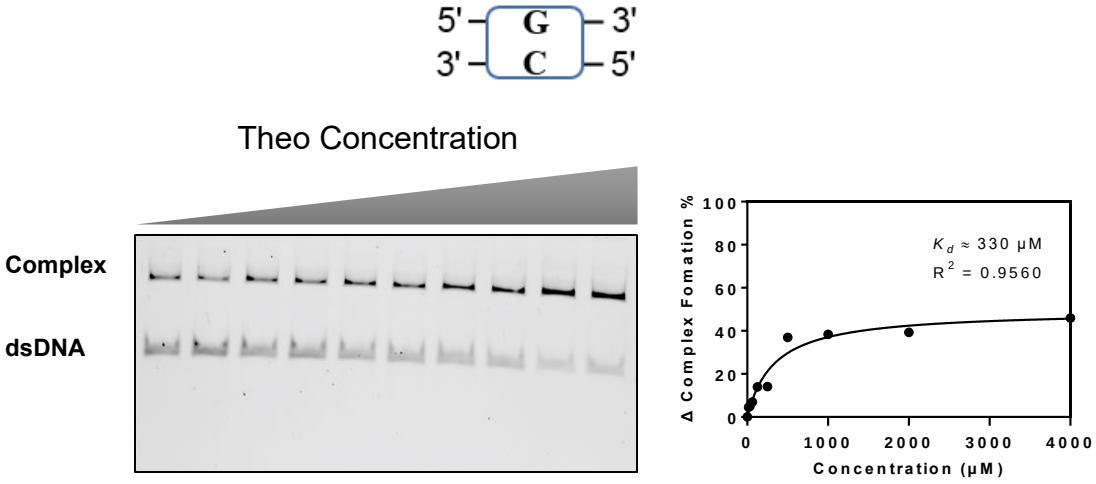

Figure S10. Concentration-dependent regulation of the complex formation with the **SL2-CM-1** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

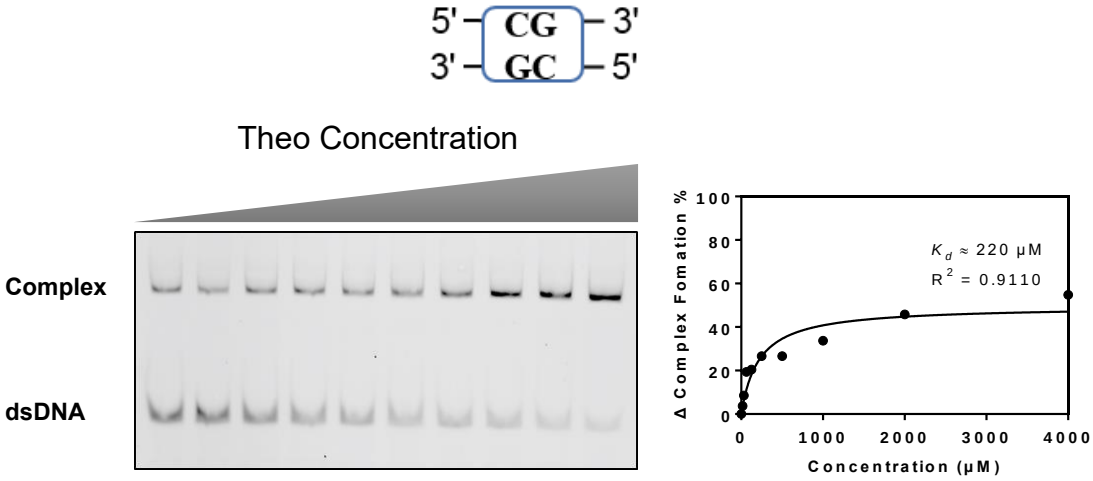

Figure S11. Concentration-dependent regulation of the complex formation with the **SL2-CM-2** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

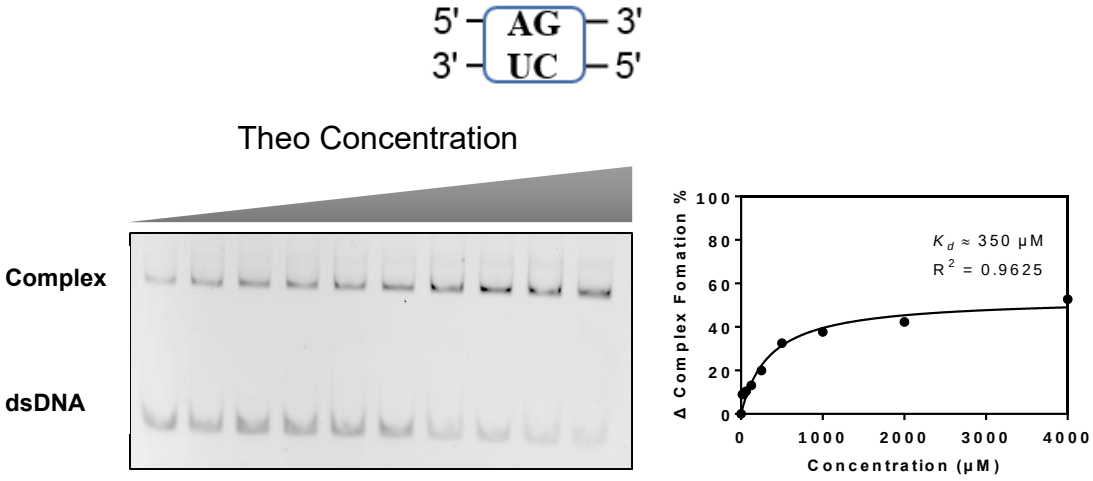

Figure S12. Concentration-dependent regulation of the complex formation with the **SL2-CM-3** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

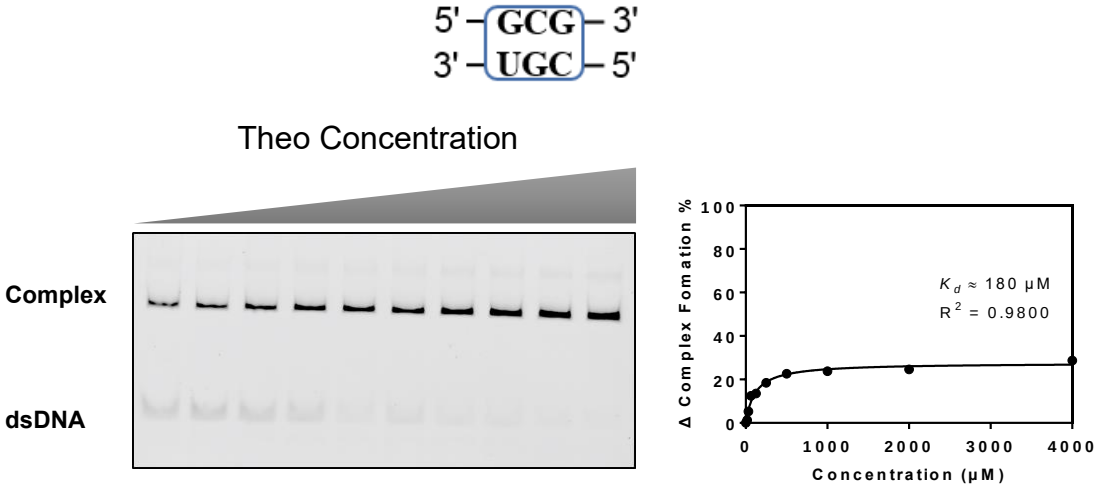

Figure S13. Concentration-dependent regulation of the complex formation with the **SL2-CM-4** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

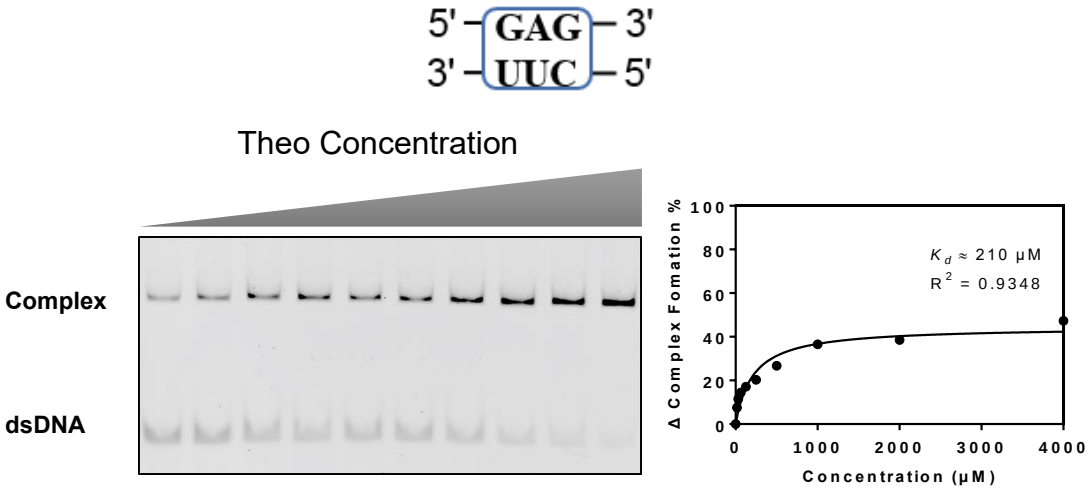

Figure S14. Concentration-dependent regulation of the complex formation with the **SL2-CM-5** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

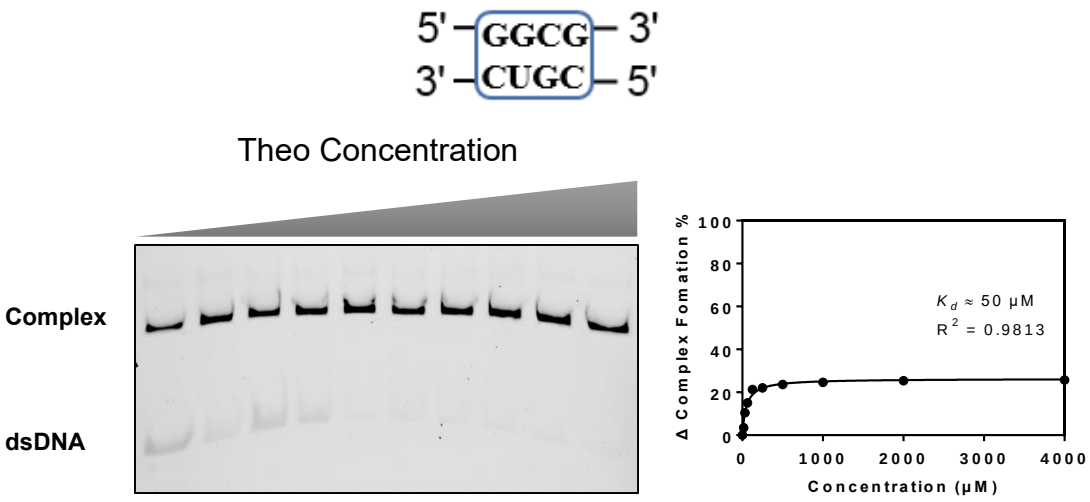

Figure S15. Concentration-dependent regulation of the complex formation with the **SL2-CM-6** design. The concentrations of theophylline were 0, 16 μM, 32 μM, 64 μM, 125 μM, 250 μM, 500 μM, 1 mM, 2 mM and 4 mM, respectively.

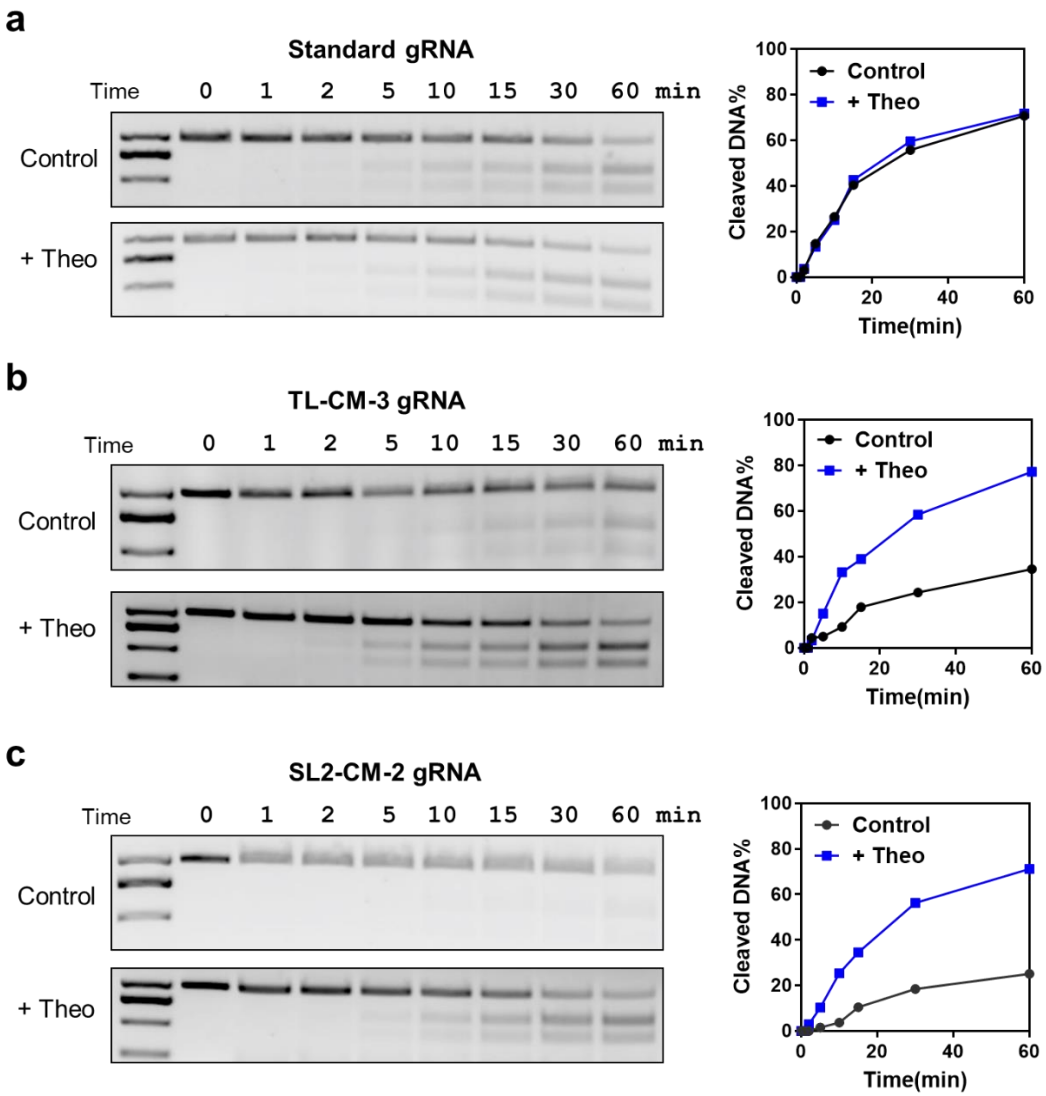

Figure S16. Cleavage assays to characterize the theophylline-controlled nuclease activity of CRISPR/Cas9 complex. Theophylline did not influence the cleavage activity of standard gRNA (**a**). In contrast, both the TL-CM-3 and SL2-CM-2 gRNAs are responsive upon the treatment of theophylline. A more efficient cleavage performance was observed in the presence of 4 mM theophylline in (**b**) and (**c**). The control samples indicated the reaction mixture without theophylline.

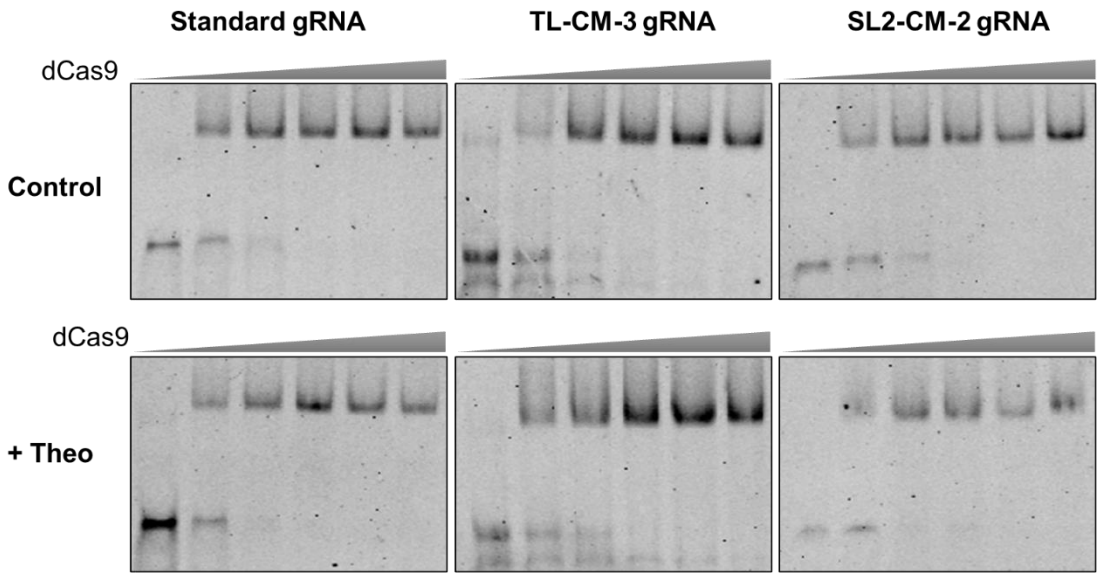

Figure S17. Binding behaviors between gRNA and the dCas9 protein. The concentration of gRNA was maintained at 20 nM; the concentrations of dCas9 were 0, 10, 20, 30, 40, and 50 nM, respectively. The binding behaviors were similar between the standard and engineered gRNAs and the treatment of theophylline had little effect on the binary complex formation. These results suggested the formation of binary complex was not significantly influenced by this approach of gRNA engineering or the treatment of theophylline.

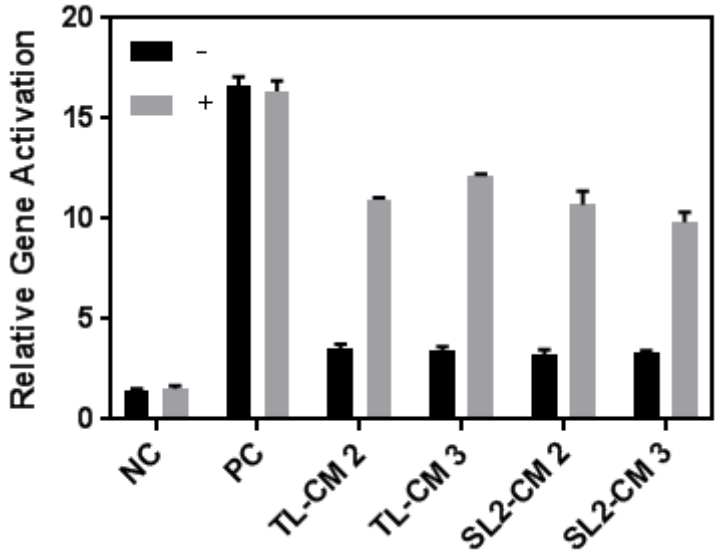

Figure S18. Relative transcriptional levels of CXCR4 gene in the presence and absence of theophylline (4 mM) in HEK293T cells. The negative control (NC) indicated the gRNA could not target the CXCR4 gene; the positive control (PC) indicated the standard gRNA without introduction of the theophylline aptamer. Engineering of the tetraloop hairpin (TL) or the stem loop 2 (SL2) with the CM-2 and CM-3 designs were examined.

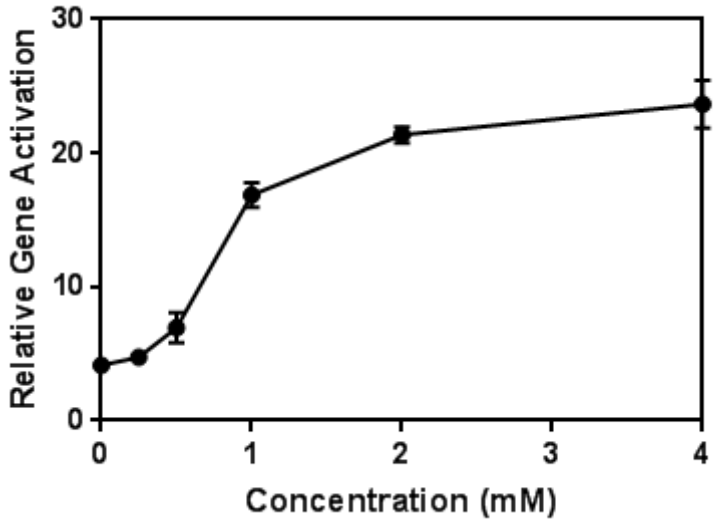

Figure S19. Concentration-dependent gene regulation of ASCL1 with the SL2-CM-2 design under the treatment of theophylline.